EngMedicine 1 (2024) 100020

Contents lists available at ScienceDirect

EngMedicine

journal homepage: www.sciencedirect.com/journal/engmedicine

Review

Artificial intelligence in orthopaedic trauma

Chuwei Tian [a,b,c,d,e], Yucheng Gao [a,b,c,d,e], Chen Rui [a,b,c,d,e], Shengbo Qin [e], Liu Shi [a,b,c,d,e], Yunfeng Rui [a,b,c,d,e,*]

[a] Department of Orthopaedics, Zhongda Hospital, School of Medicine, Southeast University, NO.87 Ding Jia Qiao, Nanjing, Jiangsu 210009, PR China
[b] Multidisciplinary Team (MDT) for Geriatric Hip Fracture Management, Zhongda Hospital, School of Medicine, Southeast University, Nanjing Jiangsu, PR China
[c] Orthopaedic Trauma Institute (OTI), Southeast University, Nanjing, Jiangsu 210009, PR China
[d] Trauma Center, Zhongda Hospital, Southeast University, Nanjing, Jiangsu 210009, PR China
[e] School of Medicine, Southeast University, NO. 87 Ding Jia Qiao, Nanjing, Jiangsu 210009, PR China

ARTICLE INFO

Keywords:
Artificial intelligence
Machine learning
Natural language processing
Orthopedic trauma
Medical decision

ABSTRACT

With the exponential growth in data processing capabilities and the progressive intertwining of medicine with industry, artificial intelligence (AI) has gained widespread application in the medical domain. Currently, AI is extensively utilized across various aspects of trauma orthopedics, including fracture identification, diagnosis and stratification, prevention strategies for falls and fractures, emergency management, and perioperative and prognostic risk assessments. This study delves into the research progress and challenges of AI in orthopedic trauma, including the clinical applications of machine learning, deep learning, and natural language processing. By illuminating these dynamic research avenues, this study aimed to catalyze interdisciplinary collaboration and spur innovation at the intersection of AI and orthopedic trauma, ultimately advancing the frontiers of patient care and clinical practice.

Introduction

The burgeoning economic growth and heightened industrialization have led to increased traumatic incidents, including traffic accidents and construction mishaps, which are salient threats to public health and are the leading cause of mortality among individuals aged <45 years [1,2]. As the intensity of diverse injury factors escalates, injury mechanisms evolve into more complex entities, complicating the management of the resultant intricate fractures. Simultaneously, the increasingly prevalent demographic shift toward an aging society has led to a steady increase in low-violence fractures, which predominantly affect older adults with preexisting health conditions and compromised systemic resilience [3]. In addressing such patient populations, healthcare providers must prioritize the restoration of daily living independence and diligently attend to their systemic health status and osteoporotic risks to prevent further harm from potential complications [4]. These distinctive attributes of trauma orthopedics significantly influence the allocation of medical resources and present novel opportunities and challenges for discipline progression.

Artificial intelligence (AI) constitutes a burgeoning interdisciplinary technical science that focuses on creating and refining theories, methodologies, technologies, and application frameworks designed to emulate, augment, and amplify human intelligence [5]. Since the 1980s, data-driven clinical prediction tools have been used to identify patients eligible for discharge [6]. The exponential growth in data processing capabilities, coupled with the progressive intertwining of medical and industrial domains, has led to the widespread adoption of AI in healthcare [7]. Substantial evidence demonstrates that AI outperforms or equals human performance in multiple tasks, such as image recognition and data analytics [8]. Relative to conventional statistical methods, AI has two key theoretical benefits: it incorporates nonlinear correlations into models and enhances model performance through self-learning [9]. Moreover, the recent surge in natural language models has prompted an increasing number of scholars to apply these advancements to clinical diagnosis and process optimization [10]. Research findings indicate that ChatGPT has promising potential to facilitate clinical documentation practices [11].

This study systematically examined the historical evolution of AI and its integration into trauma orthopedics, delving into the advancements in AI applications across various domains within trauma orthopedics, including image recognition, clinical decision-making and assessment,

* Corresponding author. Department of Orthopaedics, Zhongda Hospital, School of Medicine, Southeast University, No. 87 Ding Jia Qiao, Nanjing, Jiangsu 210009, PR China.
E-mail address: ruiyunfeng@126.com (Y. Rui).

https://doi.org/10.1016/j.engmed.2024.100020
Received 18 February 2024; Received in revised form 17 May 2024; Accepted 2 June 2024

perioperative and prognostic risk prediction, and data processing techniques. Furthermore, this review projects the potential future applications of AI in routine clinical practice, guiding future research and offering insights into the provision of precise and individualized medical solutions.

Overview of orthopedic trauma

Trauma has emerged as a paramount challenge to human health due to the increasing complexity of injury mechanisms amidst the intensifying severity of various causative factors [12,13]. Traumatic orthopedics spans a vast array of responsibilities, extending from immediate patient care to precise diagnostics, efficacious treatments, and holistic rehabilitative measures. Clinicians regularly encounter multifaceted situations characterized by intricate co-injuries, heterogeneous clinical manifestations, and variable prognoses. At its core, this necessitates that physicians possess robust data management competencies, which involve the meticulous computation and synthesis of patient injury data as well as the sophisticated understanding and integration of guidelines, expert opinions, and research evidence, constructing a coherent, logical structure linking injury severity to potential recovery paths. Despite advancements, contemporary practice continues to depend heavily on physicians' expertise for ostensibly straightforward activities, including disease classification, injury prioritization, choice of therapeutic interventions, and prognostic evaluations.

In the context of disease management, standardized medical guidelines, such as Advanced Trauma Life Support, aim to guarantee that trauma physicians base their patient treatment decisions on the most robust available evidence. Notable variations have persisted in the execution of these guidelines. A case in point is the management of hip fractures in older adults, where, despite the establishment of a relatively consensual framework, significant disputes continue to pervade aspects such as the optimal timing of intervention, choice of surgical technique, anesthetic modality, and various perioperative strategies. The continuous conducting of meta-analyses attests to the persistent nature of this debate [14,15].

The inconsistency prevalent in diagnostic and therapeutic practices highlights the fact that medicine is advancing at an unprecedented and exponential rate, becoming progressively intricate. Contemporary medical technology and expertise often find it challenging to cope with the deluge of information and accelerated tempo of change. Although expert consensus and guidelines provide indispensable guidance rooted in current evidence, they only marginally enhance clinical decision-making processes. A superior strategy involves individualized assessments and optimal management strategies according to each patient's unique characteristics.

Fundamentals of AI

AI is a comprehensive paradigm encompassing the employment of specialized mathematical algorithms that empower machines to execute problem-solving activities, engage in object and word recognition, and undertake controlled decision-making processes. Fig. 1 presents a brief history of the development of AI and representative algorithm models for various periods. This swiftly evolving and pragmatic technology addresses intricate challenges through rigorous data analysis and leveraging neural networks to emulate human cognitive capabilities [16].

Machine learning, which is a cornerstone discipline in AI, aims to converge data analysis and computational algorithms [17]. In 1959, the AI pioneer Arthur Samuel framed machine learning as a field that empowers computers to learn independently without explicit programming [18]. This conceptualization delineates three primary stages of machine learning, including data preprocessing, model training, and model evaluation. During data preprocessing, researchers define research questions, assemble datasets, and partition them into training and testing subsets. In model training, two methodologies are paramount: supervised and unsupervised learning [19]. Supervised learning involves the derivation of patterns from data by attaching accurate labels to each sample, whereas unsupervised learning facilitates algorithmic learning and pattern recognition without requiring predefined labels. Consequently, unsupervised learning typically requires larger sample sizes and more rigorous data preprocessing than supervised learning. Finally, during the evaluation phase, the efficacy of the constructed model is gauged through validation using a test set to assess its overall performance.

Recently, machine learning models have been increasingly employed in medical data processing, auxiliary diagnosis, and prognosis prediction. The decision tree is a nonparametric supervised learning algorithm that derives decision rules from datasets comprising attributes and labels and presents these rules within a tree-structured graph to address classification and regression tasks [20]. Notably, two ensemble algorithms built upon the decision tree model, random forest and gradient boosting, have demonstrated significant performance enhancements by virtue of their stacked tree structures [18]. Among the most prominent machine learning algorithms in contemporary literature, support vector machines (SVMs) stand out; they are considered among the closest precursors to deep learning techniques and have seen widespread application in image classification, data recognition, and categorization [21]. Furthermore, conventional statistical methods, such as logistic regression and naïve Bayes (NB) methods, continue to play vital roles in the research landscape of machine learning models.

Artificial neural networks (ANNs), conceptualized by McCulloch in the 1940s as machine learning algorithms that emulate human neuronal

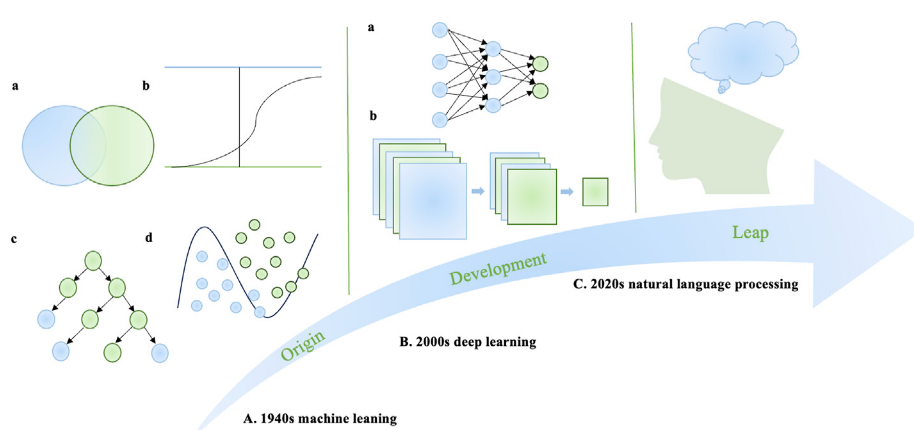

Fig. 1. A brief history of artificial intelligence. A. 1940s machine learning: a. Naïve Bayes, b. Logistic Regression, c. Decision Tree, d. Support Vector Machine. B. 2000s deep learning: a. Artificial Neural Networks, b. Convolutional Neural Networks. C. 2020s natural language processing.

processing [22], have seen their application domains grow increasingly intricate, owing to substantial advancements in computational capabilities and the exponential expansion of databases. In 2006, Hinton introduced the concept of 'deep learning,' demonstrating a method for training deep neural networks capable of recognizing handwritten digits with unprecedented accuracy [23]. These deep neural networks represent highly abstracted models of the cerebral cortex and are composed of interconnected ANNs [24]. The multilayer complexity of these networks enables algorithms to discern subtle and more complex patterns within the data compared with single- or two-layer networks. Among the prevalent deep-learning models, convolutional neural networks (CNNs) play two principal roles, including feature extraction from images and subsequent classification. CNNs initially identify elementary features within an image and then aggregate these features to derive higher-order complexities. During classification, CNNs make decisions regarding image categories and contrast them with known classifications to calculate the error margins. Through this iterative training process, CNNs can adaptively update their learning parameters and apply them to classify novel images effectively [25].

Moreover, burgeoning interest in natural language processing (NLP) tools in recent years has led to an increasing number of scholars exploring their potential in the medical domain. ChatGPT, a cutting-edge NLP instrument driven by AI technology, is a chat-generative pre-trained transformer [26]. It exhibits considerable utility in expediting clinical information retrieval, aiding clinical case design, and personalizing medicine curriculum development [27]. As algorithmic advancements continue, the ChatGPT applicability spectrum has significantly expanded.

Applications of AI in orthopedic trauma

Recently, AI has found wide application in orthopedic trauma, primarily including the following three aspects: diagnosis and imaging, treatment planning and surgical assistance, and prognosis and rehabilitation. Table 1 summarizes the representative studies included in this review.

Diagnosis and imaging

There is a strong reliance on imaging modalities, including radiography, computed tomography (CT), and magnetic resonance imaging, for the diagnosis of orthopedic diseases. Recently, rapid advancements in deep-learning algorithms have demonstrated remarkable potential for applications in the domain of image recognition, especially in the context of traumatology [25]. Notably, machine learning and derivative imaging technologies have proven to be highly effective in supporting fracture diagnosis. In the nascent stages of deep learning development, Brett et al. pioneered the application of computer-aided image recognition techniques to annotate vertebrae; they constructed a model for identifying T4–L4 vertebral fractures in non-scoliotic patients [28]. This research represents a foundational milestone in the evolution of deep learning in image recognition following the successful classification of 1.2 million high-resolution images from the ImageNet dataset by CNNs [29]. Subsequently, deep learning gained significant traction, becoming the most prominent machine learning technique widely used in the medical community.

Kim et al. trained a deep learning algorithm using 11,112 plain radiographs of distal radius fractures and normal wrist joints; subsequently, they conducted external validation to underscore the potential utility of CNNs in diagnosing distal radius fractures [30]. In instances where occult fractures prove challenging for clinicians to discern, they often resort to CT scans for differential diagnosis. However, deep learning algorithms have been shown to effectively identify such occult fractures, reducing the incidence of missed diagnoses and unnecessary examinations and conferring substantial economic benefits. Moreover, Chung et al. devised a CNN that was specifically tailored for diagnosing and neer-typing proximal humeral fractures [31]. This study demonstrated that deep learning technology is highly accurate in detecting proximal humeral fractures, with a typing accuracy of 65–85%. Significantly, when applied to complex 3- and 4-part proximal humeral fractures, deep learning algorithms outperformed orthopedic surgeons and radiologists in their assessments of plain films.

Table 1
Application of AI in orthopaedic trauma.

Application field	Year	Model type	Trauma site	Model performance	Author
Diagnosis and Imaging	2009	DL	Vertebral Fractures	Accuracy error over each vertebra was 1.06 ± 1.2 mm	Brett et al. [28]
	2018	CNN	Distal Radius Fractures	AUC 0.954	Kim et al. [30]
	2018	CNN	Proximal Humerus Fractures	Fracture recognition: Accuracy 0.96, AUC 1.0 Fracture classification: Accuracy 0.65–0.86, AUC 0.9	Chung et al. [31]
	2019	CNN	Intertrochanteric Fractures	Accuracy 0.95, Sensitivity 0.94, Specificity 0.97	Urakawa et al. [32]
	2019	CNN	Ankle Fractures	Accuracy 0.76–0.81	Kitamura et al. [34]
	2019	CNN	Calcaneus Fractures	Accuracy 0.98	Pranata et al. [35]
	2020	DL	Femoral Neck Fractures	Fracture recognition: Accuracy 0.92, AUC 0.92 Fracture classification: Accuracy 0.86, AUC 0.96	Mutasa et al. [8]
Treatment Planning and Surgical Assistance	2014	ML	Life-saving Interventions	Accuracy 0.89	Liu et al. [43]
	2020	ML	Fall and Fracture Risk	CatBoost AUC 0.688	Kong et al. [38]
	2020	ML	Fall Risk	Accuracy 0.78	Forth et al. [39]
	2020	ML	Needs-Based Assessment of Trauma Systems	Assess changes in population and injury coverage within the regions	Dooley et al. [42]
	2021	DL	Pelvic Fractures	Determine transsacral corridors S1 and S2 using artificial intelligence and 3D statistical modeling	Kamer et al. [44]
	2024	XGBoost	Distal Radius Fractures	AUC 0.74	Yamamoto et al. [40]
Prognosis and Rehabilitation	2021	ML	Hip Fractures/Mortality	AUC 0.83–0.92	DeBaun et al. [52]
	2021	ML	Hip Fractures/Mortality	AUC 0.70–0.74	Forssten et al. [53]
	2022	ML	Hip Fractures/Adverse Events	AUC 0.81–0.83	Li et al. [49]
	2023	ML	THA/Secondary Revisions	AUC 0.80–0.86	Klemt et al. [50]
	2023	ML	Hip Fractures/eLOS	AUC 0.82–0.98	Tian et al. [50]
	2023	ML	Orthopaedic Rehabilitation	R^2 0.83–0.87	Santilli et al. [55]
	2024	ML	Hip Fractures/POD	AUC 0.68–0.81	Song et al. [48]
	2024	NLP	Second Hip Fracture	AUC: 0.69 Apparent Performance: 0.75	Larrainzar-Garijo et al. [52]

Lower limb fractures are more susceptible to occult occurrences than upper limb fractures. Therefore, deep learning algorithms have significant potential as auxiliary diagnostic tools [32]. Specifically, for hip fractures, which constitute 20% of fracture cases, Urakawa et al. developed a CNN-based model for diagnosing intertrochanteric femoral fractures [33]. This study utilized a dataset consisting of 1773 fracture images and 1573 non-fracture images and demonstrated that the deep-learning model outperformed orthopedic surgeons regarding diagnostic accuracy.

Furthermore, Mutasa et al. applied a deep-learning algorithm to analyze a sample comprising 127 Garden I–II and 610 Garden II–IV hip X-rays, along with 326 normal scans, effectively identifying fractures and classifying femoral neck fractures according to the Garden classification system, achieving commendable results [8]. In the context of lower tibia-fibula and foot fractures, which are often characterized by a high rate of missed diagnoses, Kitamura et al. trained their model on an iterated dataset of 298 fracture images and 298 normal images and successfully detected concealed fractures in these regions with an accuracy of 76–81% [34]. Moreover, the application value of deep learning also extends to calcaneal fractures [35].

As research on natural language models advances, scholars have begun to generate or enhance diagnostic reports using language models while conducting imaging diagnoses with AI. Butler et al. demonstrated that foot and ankle orthopedic radiology reports, modified by AI, exhibited considerable enhancements in the Flesch reading ease score and Flesch-Kincaid reading level [36]. Future clinical practice may witness improved readability of imaging diagnosis reports through the integration of natural language models.

Previous research has established deep learning to be at least as effective as orthopedic surgeons and radiologists in image recognition tasks, particularly due to its ability to improve computational efficiency and detect occult fractures. Notably, fracture classification has a significant potential for informing treatment strategies. As the algorithm efficiency improves with an increase in the volume of available image data, the precision of deep learning for diagnosing and categorizing fractures is expected to increase incrementally. Moreover, natural language models can greatly improve the readability of imaging diagnostic reports to patients. Leveraging AI-assisted diagnosis, clinicians can expedite diagnostic and therapeutic plan formulation, optimizing the allocation of medical resources, a development poised to generate substantial economic benefits.

Treatment planning and surgical assistance

The inherent complexity of traumatic orthopedics results in significant variability in clinical decision-making processes. Recently, a growing number of researchers have incorporated machine-learning algorithms into clinical decision-making for patients with traumatic orthopedic injuries [37]. Unlike traditional logistic and linear regression techniques, many machine learning methodologies can discern nonlinear patterns, enhancing the precision of predictive outcomes [38]. Fig. 2 shows the flow of machine learning algorithm modeling and clinical transformation.

In the context of fall and fracture risk assessment, Kong et al. compared the application of the gradient boosting algorithm CatBoost, SVMs, logistic regression, and the World Health Organization Fracture Risk Assessment Tool (FRAX) for predicting fracture risk in patients with osteoporosis [39]. The study revealed that the area under the receiver operating characteristic curve (AUC) of the CatBoost algorithm significantly outperformed that of FRAX in predicting fragility fractures (0.688 vs. 0.663, $P < 0.01$). Conversely, Forth et al. utilized machine learning algorithms to analyze standing forces collected from older adults using force-measuring plates and obtained postural stability (PS) scores of 1–10 [40]. Higher PS scores denote increased stability, with an algorithm-based analysis yielding an accuracy of 0.78. Subsequent data from the authors demonstrated that seniors within the high-risk range

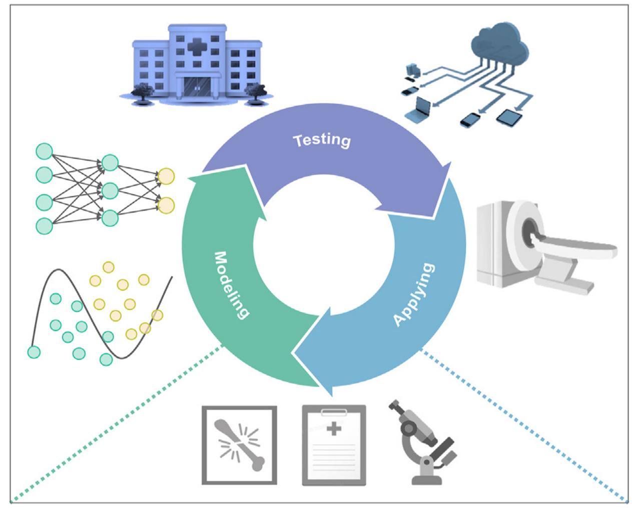

Fig. 2. Process of machine learning algorithm modeling and clinical transformation.

(PS: 1–3) were three-fold more likely to experience falls within a year than those in the low-risk category (PS: 7–10). Yamamoto et al. constructed a gait stabilization model using the XGBoost model to predict distal radius fractures and achieved satisfactory predictive accuracy [41]. Through such assessments of fall risk among the older population, one can promptly identify high-risk groups and associated risk factors, enabling a proactive preventive approach akin to 'treating without disease.'

Owing to the specialized nature of trauma orthopedics, trauma physicians often triage patients in emergency settings. When confronted with intricate cases, machine learning algorithms serve as supportive tools for enhancing the efficiency of trauma care delivery. A regional trauma system designed before patient admission optimizes the geographic placement of trauma centers and mitigates peak patient loads, reducing the strain on any individual trauma center [42]. Dooley et al. introduced a geospatially modeled trauma system evaluation tool to aid in assessing trauma incidence and center coverage across populations within regions [43]. Liu et al. devised a multiparameter machine learning algorithm and system capable of predicting the need for life-saving interventions in patients with trauma, achieving an accuracy of 89.8%, exploring the potential of machine learning to support emergency assessment and intervention for such patients [44]. In the perioperative phase, clinicians leverage machine learning models to analyze patient imaging data, facilitating optimal surgical plan formulation. Specifically, Kamer et al. employed CT-based AI three-dimensional modeling to evaluate the bone corridor of the sacroiliac joint and preoperatively planned implant positioning [45]. This technology has been applied to various procedures, including long bone fixation, joint replacement, spinal deformity correction, and post-traumatic deformity correction surgeries [46]. Technological advancements have incrementally enhanced the precision of AI-assisted preoperative planning, enabling surgeons to shorten their learning curves and execute surgeries more accurately and safely, ultimately boosting surgical success rates and minimizing surgical risks [47]. Furthermore, this technology holds promise for educating orthopedic residents, as demonstrated by Lohre et al., who found it to be significantly more efficient than conventional learning approaches [48].

Advances in the field of bone regeneration and tissue engineering have made it possible to deepen our understanding of the regeneration process by analyzing complex data using deep learning [49]. Robles-Bykbaev et al. proposed a method based on image analysis and machine learning to model and estimate the extent of bone cell regeneration and collagen degradation [50]. This study provides a novel approach to explore bone regeneration and repair and its clinical application.

Prognosis and rehabilitation

The Enhanced Recovery After Surgery (ERAS) protocol has significantly transformed perioperative care through the adoption of evidence-based practices aimed at improving patient recovery postoperatively [51]. Despite the standardized approach offered by the ERAS protocol, adjustments are necessary to account for individual patient factors such as comorbidities, age, and body composition. AI and machine learning present novel avenues for addressing these challenges and optimizing ERAS processes [52].

The integration of data analysis and predictive modeling represents one of the earliest applications of AI in the clinical context, with a multitude of studies focusing on employing machine learning to predict perioperative and long-term postoperative complication risks among patients with orthopedic trauma. Song et al. constructed a machine learning-based prediction model for postoperative delirium in patients with hip fractures that included a predictive nomogram with good accuracy [53]. The key risk factors identified for postoperative delirium in older patients with hip fractures were age, renal insufficiency, chronic obstructive pulmonary disease, antipsychotic usage, lactate dehydrogenase level, and C-reactive protein level. Li et al. developed another machine learning-driven predictive model to assess the likelihood of postoperative complications, intensive care unit (ICU) transfer, and delayed discharge in patients with hip fractures and showed that this model significantly outperformed the American Society of Anesthesiologists Physical Status score with superior predictive power (postoperative complications: 0.810 vs. 0.629, $P < 0.01$; ICU transfer: 0.835 vs. 0.692, $P < 0.01$; delayed discharge: 0.832 vs. 0.618, $P < 0.01$) [54]. Our previous research also illustrated analogous outcomes, demonstrating that machine learning outperforms conventional statistical models (AUC: 0.82–0.98) in predicting extended lengths of stays among geriatric patients with hip fractures and in ascertaining associated risk factors [55]. Regarding long-term postoperative complications, Klemt et al. applied machine learning algorithms to ascertain the risk factors for secondary revisions in patients who underwent total hip replacement; they revealed that the four implemented models demonstrated outstanding discriminative capacity, calibration, and decision-curve performance [56].

Furthermore, Larrainzar-Garijo et al. innovatively used NLP to mine electronic medical records and extract relevant attributes to forecast the risk of recurrent hip fractures, achieving promising results [57]. This study leveraged the rapid advancements in AI within NLP, presenting new perspectives for future AI research endeavors. Leveraging AI methods to identify perioperative complications and associated risk factors, clinicians can expeditiously identify high-risk patients and manage them proactively, providing a pioneering avenue for refining perioperative diagnosis and treatment protocols and mitigating complication rates.

Machine learning algorithms have demonstrated notable efficacy in predicting patient prognosis and recovery. DeBaun et al. investigated the comparative performance of various algorithms in predicting 30-day mortality rates after hip fractures and reported that ANNs achieved the highest AUC of 0.92, followed closely by logistic regression (AUC: 0.87) and NB regression (AUC: 0.83) [58]. In a separate study, Forssten et al. juxtaposed the predictive powers of SVMs, NB, random forest, and logistic regression algorithms regarding 1-year postoperative mortality in patients with hip fractures, indicating that SVMs, NB, and random forest may lead to an overestimation of patient mortality [59]. In contrast, logistic regression emerged as the most reliable method among the four methods for predicting long-term outcomes. This research underscores the enduring utility of traditional statistical methods in prognosis prediction while highlighting that machine learning algorithms require a heightened level of data homogeneity within large datasets to realize their predictive capacity fully. In addition, given the critical state of patients with orthopedic trauma in ICU settings, early identification of patients at a high risk of death has important clinical implications. Han et al. developed and validated an AI model for predicting the 30-day mortality in patients with critical fracture trauma and deployed it online, allowing physicians to obtain the predictive risk of death for critically ill patients by inputting patient characteristics [60]. The eXGBM algorithm showed the highest predictive performance, and external verification revealed that the AUC of the model developed by the authors was 0.913.

Santilli et al. developed four integrated decision-tree-based models aimed at assessing the rehabilitative potential of post-discharge programs for patients with orthopedic and neurological disorders intended to improve their functional abilities [61]. Through the meticulous analysis and interpretation of patient rehabilitation data, these models can accurately predict rehabilitation progression, enabling physicians to make prompt adjustments to maximize rehabilitation results. As the volume of available clinical patient data continues to expand and personalized medicine advances, machine learning algorithms are being developed to gain a broader application scope.

Challenges and considerations

Progressive advancements in science and technology, coupled with enhanced algorithmic efficiency, have led to a steady increase in learning applications within the medical sector [62]. Machine learning algorithms are increasingly employed as ancillary instruments in orthopedic trauma, revealing substantial potential across various domains, such as fracture identification and categorization, perioperative care management, and patient outcome prediction [9,63]. However, certain obstacles hinder the implementation of these algorithms.

1. Ethical considerations in machine learning research and applications are paramount, particularly concerning the assurance of data accuracy and privacy preservation throughout the process [64]. Despite the absence of a universally accepted ethical framework in the machine learning domain, adherence to ethical principles must remain a cornerstone of research endeavors.
2. The quality and volume of data in clinical databases play pivotal roles, as machine learning models require extensive, high-quality information. The intricacy of characterizing orthopedic trauma cases complicates the process of guaranteeing homogeneity and diversity within assembled datasets. Consequently, efforts must be directed toward significantly expanding the scale of clinical samples and enriching their feature profiles to furnish a model with sufficient and diverse data for robust fitting.
3. The interpretability of machine learning models constitutes a critical challenge despite their superior data-fitting capabilities relative to conventional statistical techniques. The inherent tendency of machine learning models, especially complex models such as deep neural networks, to avoid overfitting renders them somewhat enigmatic or 'black boxes,' making it difficult to elucidate their output and decision-making rationale [65]. Consequently, developing methodologies for evaluating the interpretability of machine learning models and devising inherently interpretable machine learning applications are pressing concerns that require immediate attention.
4. The limitations of universal models in accommodating personalized care requirements are noteworthy, as the clinical management of orthopedic trauma necessitates tailored treatment strategies. Clinical medicine plays a pivotal role in orthopedic trauma, given that it transcends procedural determinism, discretionary judgment, and targeted interventions by clinicians. Considering the unique characteristics of each patient's injury type, severity, and physiological state, standard machine-learning models may not adequately address individualized needs. Consequently, the development of bespoke algorithms and patient-specific treatment protocols is promising for future research.

Future trends and possibilities

Despite its inherent limitations, AI holds substantial promise for applications in trauma orthopedic clinical practice. A crucial aspect of this approach is the accurate classification of trauma types to devise effective treatment strategies. Leveraging deep learning models that process extensive clinical and imaging datasets, AI enables the precise identification and categorization of various fracture types, enhancing physicians' diagnostic precision and therapeutic decision-making. This capability contributes to reduced diagnostic latency, minimized misdiagnosis rates, and improved patient recovery outcomes. Furthermore, through the analysis of individual patient clinical records, physiological metrics, and imaging data, AI-driven machine-learning models can aid in forecasting the likelihood of trauma incidents. Such predictive insights will empower healthcare providers to implement timely and targeted interventions, ultimately reducing the incidence of trauma in at-risk individuals. These advancements have profound implications for the surveillance and management of high-risk populations. Finally, the integration of NLP algorithms introduces untapped potential for AI in domains such as clinical pedagogy and case administration.

The anticipated advancements in AI research hold promise in making clinical practices in trauma orthopedics increasingly precise and individualized. The integration of methods such as reinforcement and transfer learning has the potential to enhance model performance and generalizability significantly. Concurrently, fostering multicenter collaborations through data-sharing initiatives and constructing extensive clinical databases are pivotal to mitigating the challenges of limited data resources. Moreover, while propelling the application of machine learning within trauma orthopedics, it is imperative to meticulously contemplate various elements, including data privacy, model transparency, and clinical practicality, by constantly refining algorithms and their applications to achieve optimal clinical results. Recently, the introduction of guidelines for implementing AI in orthopedic research has been welcomed [66,67]. This introduction of guidelines for experimental research and clinical applications can enhance the standardization of AI utilization among researchers, addressing ethical concerns arising from its misuse.

Conclusion

The recent rapid advancements in technology and algorithm efficiency have significantly broadened the application scope of AI in the domain of traumatic orthopedics. AI algorithms exhibit good precision in image data modeling, encompassing fracture identification, diagnosis, classification, and various aspects of clinical decision-making, such as fall and fracture prevention strategies, emergency response measures, perioperative risk assessment, and prognostic forecasting. Currently, the implementation of AI in trauma orthopedics is in an exploratory stage. However, with ongoing scientific and technological developments and an exponential increase in available data, machine learning and deep learning models are poised for substantial improvements in accuracy and practical utility. These models have promising potential applications in areas such as diagnostic assistance, decision support, personalization of medical interventions, and efficient allocation of clinical resources.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

CWT designed this study and wrote the manuscript. YCG, CR and SBQ contributed to the writing of the manuscript. LS and YFR provided technical support. YFR provided the idea and revised and proofread the paper. All the authors read and approved the final manuscript.

CRediT authorship contribution statement

Chuwei Tian: Writing – review & editing, Writing – original draft, Conceptualization. **Yucheng Gao:** Writing – review & editing, Writing – original draft, Conceptualization. **Chen Rui:** Writing – review & editing. **Shengbo Qin:** Data curation. **Liu Shi:** Writing – review & editing, Conceptualization. **Yunfeng Rui:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank all the participants included in this current study.

References

[1] Dauer E, Goldberg A. What's new in trauma resuscitation? Adv Surg 2019. https://doi.org/10.1016/j.yasu.2019.04.010.
[2] Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012. https://doi.org/10.1016/S0140-6736(12)61728-0.
[3] Clare D, Zink KL. Geriatric trauma. Emerg Med Clin 2021. https://doi.org/10.1016/j.emc.2021.01.002.
[4] Feng JN, Zhang CG, Li BH, Zhan SY, Wang SF, Song CL. Global burden of hip fracture: the global burden of disease study. Osteoporos Int 2024. https://doi.org/10.1007/s00198-023-06907-3.
[5] Myers TG, Ramkumar PN, Ricciardi BF, Urish KL, Kipper J, Ketonis C. Artificial intelligence and orthopaedics: an introduction for clinicians. J Bone Joint Surg Am 2020. https://doi.org/10.2106/JBJS.19.01128.
[6] Goldman LCE, Johnson Pa, da Brand, rouan gW, lee tH. Prediction of the need for intensive care in patients who come to emergency departments with acute chest pain. N Engl J Med 1996. https://doi.org/10.1056/NEJM199606063342303.
[7] Paixao GMM, Santos BC, Araujo RM, Ribeiro MH, Moraes JL, Ribeiro AL. Machine learning in medicine: review and applicability. Arq Bras Cardiol 2022. https://doi.org/10.36660/abc.20200596.
[8] Mutasa S, Varada S, Goel A, Wong TT, Rasiej MJ. Advanced deep learning techniques applied to automated femoral neck fracture detection and classification. J Digit Imag 2020. https://doi.org/10.1007/s10278-020-00364-8.
[9] Oosterhoff JHF, Doornberg JN, Machine Learning C. Artificial intelligence in orthopaedics: false hope or not? A narrative review along the line of Gartner's hype cycle. EFORT Open Rev 2020. https://doi.org/10.1302/2058-5241.5.190092.
[10] Agharia S, Szatkowski J, Fraval A, Stevens J, Zhou Y. The ability of artificial intelligence tools to formulate orthopaedic clinical decisions in comparison to human clinicians: an analysis of ChatGPT 3.5, ChatGPT 4, and Bard. J Orthop 2024. https://doi.org/10.1016/j.jor.2023.11.063.
[11] Baker HP, Dwyer E, Kalidoss S, Hynes K, Wolf J, Strelzow JA. ChatGPT's ability to assist with clinical documentation: a randomized controlled trial. J Am Acad Orthop Surg 2024. https://doi.org/10.5435/JAAOS-D-23-00474.
[12] Veronese N, Maggi S. Epidemiology and social costs of hip fracture. Injury 2018. https://doi.org/10.1016/j.injury.2018.04.015.
[13] Bhushan S, Huang X, Duan Y, Xiao Z. The impact of regional versus general anesthesia on postoperative neurocognitive outcomes in elderly patients undergoing hip fracture surgery: a systematic review and meta-analysis. Int J Surg 2022. https://doi.org/10.1016/j.ijsu.2022.106854.
[14] Park C, Bharija A, Mesias M, Mitchell A, Krishna P, Storr-Street N, et al. Association between implementation of a geriatric trauma clinical pathway and changes in rates of delirium in older adults with traumatic injury. JAMA Surg 2022. https://doi.org/10.1001/jamasurg.2022.1556.
[15] Van Waesberghe J, Stevanovic A, Rossaint R, Coburn M. General vs. neuraxial anaesthesia in hip fracture patients: a systematic review and meta-analysis. BMC Anesthesiol 2017. https://doi.org/10.1186/s12871-017-0380-9.
[16] Sardar P, Abbott JD, Kundu A, Aronow HD, Granada JF, Giri J. Impact of artificial intelligence on interventional cardiology: from decision-making aid to advanced interventional procedure assistance. JACC Cardiovasc Interv 2019. https://doi.org/10.1016/j.jcin.2019.04.048.
[17] Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med 2016. https://doi.org/10.1056/NEJMp1606181.
[18] Deo RC. Machine learning in medicine. Circulation 2015. https://doi.org/10.1161/CIRCULATIONAHA.115.001593.
[19] Barreto GA, Souza LG. Adaptive filtering with the self-organizing map: a performance comparison. Neural Network 2006. https://doi.org/10.1016/j.neunet.2006.05.005.
[20] Podgorelec VKPSB, Rozman I. Decision trees: an overview and their use in medicine. J Med Syst 2002. https://doi.org/10.1023/a:1016409317640.

[21] Mavroforakis ME, Theodoridis S. A geometric approach to support vector machine (SVM) classification. IEEE Trans Neural Network 2006. https://doi.org/10.1109/TNN.2006.873281.
[22] PW MW. A logical calculus of the ideas immanent in nervous activity. Bull Math Biol 1943. https://doi.org/10.1007/bf02478259.
[23] Geoffrey E, Hinton SOY-WT. A fast learning algorithm for deep belief nets. Neural Comput 2006. https://doi.org/10.1162/neco.2006.18.7.1527.
[24] Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg 2018. https://doi.org/10.1097/SLA.0000000000002693.
[25] Lalehzarian SP, Gowd AK, Liu JN. Machine learning in orthopaedic surgery. World J Orthoped 2021. https://doi.org/10.5312/wjo.v12.i9.685.
[26] Gordijn B, Have HT. ChatGPT: evolution or revolution? Med Health Care Philos 2023. https://doi.org/10.1007/s11019-023-10136-0.
[27] Leng L. Challenge, integration, and change: ChatGPT and future anatomical education. Med Educ Online 2024. https://doi.org/10.1080/10872981.2024.2304973.
[28] Alan Brett CGM, Hayes Curtis W, Krasnow Joel, Ozanian Takouhi, Abrams Ken, Block Jon E, et al. Development of a clinical workflow tool to enhance the detection of vertebral fractures. Spine 2009. https://doi.org/10.1097/BRS.0b013e3181b2eb69.
[29] Krizhevsky ASI, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017. https://doi.org/10.1145/3065386.
[30] Kim DH, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clin Radiol 2018. https://doi.org/10.1016/j.crad.2017.11.015.
[31] Chung SW, Han SS, Lee JW, Oh KS, Kim NR, Yoon JP, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. Acta Orthop 2018. https://doi.org/10.1080/17453674.2018.1453714.
[32] Meena T, Roy S. Bone fracture detection using deep supervised learning from radiological images: a paradigm shift. Diagnostics 2022. https://doi.org/10.3390/diagnostics12102420.
[33] Urakawa T, Tanaka Y, Goto S, Matsuzawa H, Watanabe K, Endo N. Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. Skeletal Radiol 2019. https://doi.org/10.1007/s00256-018-3016-3.
[34] Kitamura G, Chung CY, Moore 2nd BE. Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation. J Digit Imag 2019. https://doi.org/10.1007/s10278-018-0167-7.
[35] Pranata YD, Wang KC, Wang JC, Idram I, Lai JY, Liu JW, et al. Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images. Comput Methods Progr Biomed 2019. https://doi.org/10.1016/j.cmpb.2019.02.006.
[36] Butler JJ, Harrington MC, Tong Y, Rosenbaum AJ, Samsonov AP, Walls RJ, et al. From jargon to clarity: improving the readability of foot and ankle radiology reports with an artificial intelligence large language model. Foot Ankle Surg 2024. https://doi.org/10.1016/j.fas.2024.01.008.
[37] Poduval M, Ghose A, Manchanda S, Bagaria V, Sinha A. Artificial intelligence and machine learning: a new disruptive force in orthopaedics. Indian J Orthop 2020. https://doi.org/10.1007/s43465-019-00023-3.
[38] Lee NJ, Kothari P, Phan K, Shin JI, Cutler HS, Lakomkin N, et al. Incidence and risk factors for 30-day unplanned readmissions after elective posterior lumbar fusion. Spine 2018. https://doi.org/10.1097/BRS.0000000000001586.
[39] Kong SH, Ahn D, Kim BR, Srinivasan K, Ram S, Kim H, et al. A novel fracture prediction model using machine learning in a community-based cohort. JBMR Plus 2020. https://doi.org/10.1002/jbm4.10337.
[40] Forth KE, Wirfel KL, Adams SD, Rianon NJ, Lieberman Aiden E, Madansingh SI. A postural assessment utilizing machine learning prospectively identifies older adults at a high risk of falling. Front Med 2020. https://doi.org/10.3389/fmed.2020.591517.
[41] Yamamoto A, Fujita K, Yamada E, Ibara T, Nihey F, Inai T, et al. Gait characteristics in patients with distal radius fracture using an in-shoe inertial measurement system at various gait speeds. Gait Posture 2024. https://doi.org/10.1016/j.gaitpost.2023.10.023.
[42] Stonko DP, Guillamondegui OD, Fischer PE, Dennis BM. Artificial intelligence in trauma systems. Surgery 2021. https://doi.org/10.1016/j.surg.2020.07.038.
[43] Dooley JH, Ozdenerol E, Sharpe JP, Magnotti LJ, Croce MA, Fischer PE. Location, location, location: utilizing needs-based assessment of trauma systems-2 in trauma system planning. J Trauma Acute Care Surg 2020. https://doi.org/10.1097/TA.0000000000002463.
[44] Liu NT, Holcomb JB, Wade CE, Batchinsky AI, Cancio LC, Darrah MI, et al. Development and validation of a machine learning algorithm and hybrid system to predict the need for life-saving interventions in trauma patients. Med Biol Eng Comput 2014. https://doi.org/10.1007/s11517-013-1130-x.
[45] Kamer L, Noser H, Arand C, Handrich K, Rommens PM, Wagner D. Artificial intelligence and CT-based 3D statistical modeling to assess transsacral corridors and plan implant positioning. J Orthop Res 2021. https://doi.org/10.1002/jor.25010.
[46] Fadero PE, Shah M. Three dimensional (3D) modelling and surgical planning in trauma and orthopaedics. Surgeon 2014. https://doi.org/10.1016/j.surge.2014.03.008.
[47] Youssef Y, De Wet D, Back DA, Scherer J. Digitalization in orthopaedics: a narrative review. Front Surg 2023. https://doi.org/10.3389/fsurg.2023.1325423.
[48] Lohre R, Bois AJ, Athwal GS, Goel DP, Canadian S, Elbow S. Improved complex skill acquisition by immersive virtual reality training: a randomized controlled trial. J Bone Joint Surg Am 2020. https://doi.org/10.2106/JBJS.19.00982.
[49] Jeyaraman M, Ratna HVK, Jeyaraman N, Venkatesan A, Ramasubramanian S, Yadav S. Leveraging artificial intelligence and machine learning in regenerative orthopedics: a paradigm shift in patient care. Cureus 2023. https://doi.org/10.7759/cureus.49756.
[50] Robles-Bykbaev Y, Naya S, Diaz-Prado S, Calle-Lopez D, Robles-Bykbaev V, Garzon L, et al. An artificial-vision- and statistical-learning-based method for studying the biodegradation of type I collagen scaffolds in bone regeneration systems. PeerJ 2019. https://doi.org/10.7717/peerj.7233.
[51] Mithany RH, Daniel N, Shahid MH, Aslam S, Abdelmaseeh M, Gerges F, et al. Revolutionizing surgical care: the power of enhanced recovery after surgery (ERAS). Cureus 2023. https://doi.org/10.7759/cureus.48795.
[52] Zain Z, Almadhoun M, Alsadoun L, Bokhari SFH. Leveraging artificial intelligence and machine learning to optimize enhanced recovery after surgery (ERAS) protocols. Cureus 2024. https://doi.org/10.7759/cureus.56668.
[53] Song Y, Zhang D, Wang Q, Liu Y, Chen K, Sun J, et al. Prediction models for postoperative delirium in elderly patients with machine-learning algorithms and SHapley Additive exPlanations. Transl Psychiatry 2024. https://doi.org/10.1038/s41398-024-02762-w.
[54] Li YY, Wang JJ, Huang SH, Kuo CL, Chen JY, Liu CF, et al. Implementation of a machine learning application in preoperative risk assessment for hip repair surgery. BMC Anesthesiol 2022. https://doi.org/10.1186/s12871-022-01648-y.
[55] Tian CW, Chen XX, Shi L, Zhu HY, Dai GC, Chen H, et al. Machine learning applications for the prediction of extended length of stay in geriatric hip fracture patients. World J Orthoped 2023. https://doi.org/10.5312/wjo.v14.i10.741.
[56] Klemt C, Cohen-Levy WB, Robinson MG, Burns JC, Alpaugh K, Yeo I, et al. Can machine learning models predict failure of revision total hip arthroplasty? Arch Orthop Trauma Surg 2023. https://doi.org/10.1007/s00402-022-04453-x.
[57] Larrainzar-Garijo R, Fernandez-Tormos E, Collado-Escudero CA, Alcantud Ibanez M, Onorbe-San Francisco F, Marin-Corral J, et al. Predictive model for a second hip fracture occurrence using natural language processing and machine learning on electronic health records. Sci Rep 2024. https://doi.org/10.1038/s41598-023-50762-5.
[58] DeBaun MR, Chavez G, Fithian A, Oladeji K, Van Rysselberghe N, Goodnough LH, et al. Artificial neural networks predict 30-day mortality after hip fracture: insights from machine learning. J Am Acad Orthop Surg 2021. https://doi.org/10.5435/JAAOS-D-20-00429.
[59] Forssten MP, Bass GA, Ismail AM, Mohseni S, Cao Y. Predicting 1-year mortality after hip fracture surgery: an evaluation of multiple machine learning approaches. J Personalized Med 2021. https://doi.org/10.3390/jpm11080727.
[60] Han T, Xiong F, Sun B, Zhong L, Han Z, Lei M. Development and validation of an artificial intelligence mobile application for predicting 30-day mortality in critically ill patients with orthopaedic trauma. Int J Med Inf 2024. https://doi.org/10.1016/j.ijmedinf.2024.105383.
[61] Santilli V, Mangone M, Diko A, Alviti F, Bernetti A, Agostini F, et al. The use of machine learning for inferencing the effectiveness of a rehabilitation program for orthopedic and neurological patients. Int J Environ Res Publ Health 2023. https://doi.org/10.3390/ijerph20085575.
[62] Hendrickx LAM, Sobol GL, Langerhuizen DWG, Bulstra AEJ, Hreha J, Sprague S, et al. A machine learning algorithm to predict the probability of (occult) posterior malleolar fractures associated with tibial shaft fractures to guide "malleolus first" fixation. J Orthop Trauma 2020. https://doi.org/10.1097/BOT.0000000000001663.
[63] Kruse C, Eiken P, Vestergaard P. Machine learning principles can improve hip fracture prediction. Calcif Tissue Int 2017. https://doi.org/10.1007/s00223-017-0238-7.
[64] Chernew MELM. Targeted supplemental data collection - addressing the quality-measurement conundrum. N Engl J Med 2018. https://doi.org/10.1056/NEJMp1713834.
[65] Halevy ANPPF. The unreasonable effectiveness of data. IEEE Intell Syst 2009. https://doi.org/10.1109/MIS.2009.36.
[66] Zsidai B, Hilkert AS, Kaarre J, Narup E, Senorski EH, Grassi A, et al. A practical guide to the implementation of AI in orthopaedic research - part 1: opportunities in clinical application and overcoming existing challenges. J Exp Orthop 2023. https://doi.org/10.1186/s40634-023-00683-z.
[67] Zsidai B, Kaarre J, Narup E, Hamrin Senorski E, Pareek A, Grassi A, et al. A practical guide to the implementation of artificial intelligence in orthopaedic research-Part 2: a technical introduction. J Exp Orthop 2024. https://doi.org/10.1002/jeo2.12025.